HU-EP-09/04

SFB/CCP-09-20

Precision check on triviality of ϕ^4 theory by a new simulation method

Ulli Wolff[1]

[1]*Institut für Physik, Humboldt Universität, Newtonstr. 15, 12489 Berlin, Germany*

We report precise simulations of ϕ^4 theory in the Ising limit. A recent technique to stochastically evaluate the all-order strong coupling expansion is combined with exact identities in the closely related Aizenman random current representation. In this way estimates of the renormalized coupling close to the continuum limit become possible with unprecedented precision and yet low CPU cost. As a sample application we present results for the unbroken phase of the Ising model in dimensions 3, 4 and 5 and investigate the question of triviality by studying a finite size scaling continuum limit.

The ϕ^4 theory of a real scalar field is the starting point of many textbooks on quantum field theory. It also plays a phenomenological rôle as an extremely simplified model of the Higgs sector of the standard model. In [1], building on [2], it was discussed that the theory is trivial in dimension four, meaning that no interaction can exist in the continuum limit. For applications this means, that the effective theory *with* an unremovable cutoff in place has only a limited energy domain of validity. This important property is rigorously known to hold above four dimensions [3], [4], but in $D = 4$ we have to rely so-far on numerical checks. In this letter we report on the discovery of a new numerical strategy and algorithm, which enhances the precision of such checks by orders of magnitude as the continuum limit is approached. Also novel is the use of a finite volume renormalization scheme in this context. Due to both improvements we can progress much deeper into the universal scaling region than in previous computations like [5]. We find a 'borderline' agreement with standard perturbation theory for the cutoff dependence of the interaction strength which calls for

more study.

We here consider the Ising limit of ϕ^4 on a D-dimensional hypercubic lattice of extent L and lattice spacing a in all directions. We mostly use lattice units $a = 1$ from here on, but occasionally re-introduce a to emphasize cutoff dependencies. Physical information is extracted via n-point correlation functions

$$\langle s(x_1)s(x_2)\cdots s\,(x_n)\rangle = \frac{Z(x_1, x_2, \ldots, x_n)}{Z(\emptyset)} \tag{1}$$

with

$$Z(x_1, x_2, \ldots, x_n) = 2^{-V} \sum_s \mathrm{e}^{\beta \sum_{l=\langle xy\rangle} s(x)s(y)} s(x_1)s(x_2)\cdots s\,(x_n) \tag{2}$$

and the volume $V = L^D$. We sum over all Ising configurations $s(x) = \pm 1$ and $Z(\emptyset)$ is the proper partition function with no field insertions. On our finite lattice $Z(.)$ is analytic for all values of β. We parametrize the strong coupling expansion in β by summing in addition to s over an integer link field $k(l) = 0, 1, \ldots, \infty$

$$Z(x_1, x_2, \ldots, x_n) = 2^{-V} \sum_{s,k} w[k] \prod_{l=\langle xy\rangle} [s(x)s(y)]^{k(l)} s(x_1)s(x_2)\cdots s\,(x_n) \tag{3}$$

with the multiple Poisson weight

$$w[k] = \prod_l \frac{\beta^{k(l)}}{k(l)!}. \tag{4}$$

For each k the s sum can now be performed and leaves behind a constraint for k,

$$Z(x_1, x_2, \ldots, x_n) = \sum_k w[k]\delta_{Q[k],X}. \tag{5}$$

The Kronecker δ enforces the coincidence of two sets. The source set $Q[k]$ consists of the sites surrounded by an *odd* total number of $k(l)$,

$$Q[k] = \left\{ x | \sum_{l,\partial l \ni x} k(l) = 1 (\mathrm{mod}\, 2) \right\}. \tag{6}$$

The insertion set X coincides with $(x_1, x_2, \ldots, x_n)$ if they are mutually different, but is more generally given by

$$X = \left\{ x | \sum_{i=1}^{n} \delta_{x,x_i} = 1 (\mathrm{mod}\, 2) \right\}. \tag{7}$$

Michael Aizenman [3], [6] has used the above representation of the Ising model to obtain rigorous correlation inequalities. He calls $\{k(l)\}$ random currents and the sets $Q = X$ defects

or external sources. Among his results the following is of interest here as it can be turned into an efficient numerical algorithm. From proposition 5.1 in [6] the identity

$$Z_c(x_1, x_2, x_3, x_4) = -2 \sum_{k,k'} w[k] w[k'] \delta_{Q[k],X_{12}} \delta_{Q[k'],X_{34}} \mathcal{X}(x_1, x_3; k + k') \tag{8}$$

with the connected part of Z_c of Z

$$\begin{aligned} Z_c(x_1, x_2, x_3, x_4) &= Z(x_1, x_2, x_3, x_4) - Z(x_1, x_2) Z(x_3, x_4) \\ &- Z(x_1, x_3) Z(x_2, x_4) - Z(x_1, x_4) Z(x_2, x_3) \end{aligned} \tag{9}$$

follows. The sets X_{12} and X_{34} are formed as in (7) but from only two points each. The cluster incidence function $\mathcal{X} \in \{0, 1\}$ is one if x_1, x_3 are in the same bond percolation cluster built by bonds that are active on links where $k(l) + k'(l) > 0$ holds in the doubled random current system. Note that the pairs $\{x_1, x_2\}$ and $\{x_3, x_4\}$ are connected automatically for k, k' that contribute.

In the Monte Carlo community is has recently been found [7], [8], [9] that it is both possible and advantageous to simulate the untruncated strong coupling expansion instead of the original path integral (or sum) over fields. In a simple variant of the worm algorithm [7] one simulates the ensemble corresponding to the partition function

$$\mathcal{Z} = \sum_{u,v,k} w[k] \delta_{Q[k],X_{uv}} \tag{10}$$

with a corresponding definition of expectation values of observables $\langle\langle \mathcal{O}[u, v, k] \rangle\rangle$. In the sum u, v run over all lattice sites.

To simulate this ensemble our elementary update step is as follows:

- pick at random one of the $2D$ links emanating from u and call it $l = \langle u\tilde{u} \rangle$
- assign a new value $\tilde{k}(l)$ to this link with probability $p_{\tilde{k}} = \exp(-\beta)\beta^{\tilde{k}}/\tilde{k}!$
- if $\tilde{k} - k$ is odd, move $u \to \tilde{u}$, otherwise leave u unchanged

If we alternate these steps with similar ones for v we have a correct algorithm for (10). Ergodicity may be shown by steps deforming an arbitrary configuration to the trivial one. This local heatbath has proved to be slightly superior to Metropolis proposals with $k(l) \to$

$k(l) \pm 1$. It is not difficult to show that the Ising two-point function is now given by the ratio of histograms

$$\langle s(x)s(0)\rangle = \frac{\langle\langle \delta_{x,u-v}\rangle\rangle}{\langle\langle \delta_{u,v}\rangle\rangle} \tag{11}$$

where we have used translation invariance. This implies in particular that the susceptibility is given by

$$\chi_2 = \sum_x \langle s(x)s(0)\rangle = [\langle\langle \delta_{u,v}\rangle\rangle]^{-1}. \tag{12}$$

To now make use of (8) for the connected four point susceptibility

$$\chi_4 = \sum_{x,y,z} \langle s(x)s(y)s(z)s(0)\rangle_c \tag{13}$$

with subtractions as in (9) all we have to do is simulate two independent replica of (10) and sum over all x_i in (8) to arrive at

$$-V\chi_4 = 2\frac{\langle\langle \mathcal{X}(u,u',k+k')\rangle\rangle}{\langle\langle \delta_{u,v}\delta_{u',v'}\rangle\rangle}. \tag{14}$$

In total we thus have derived

$$-\frac{1}{V}\frac{\chi_4}{(\chi_2)^2} = 2\left\langle\langle \mathcal{X}(u,u',k+k')\rangle\right\rangle. \tag{15}$$

The right hand side is obviously bounded between 0 and 2. In particular the lower bound corresponds to the Lebowitz inequality. Our estimator reflects this property manifestly, the subtraction of disconnected parts has been achieved analytically. We expect this to lead to a superior precision for χ_4 compared to conventional Monte Carlo procedures since they involve substantial numerical cancellations here with the correspondingly enhanced relative errors.

The previous expression is strongly reminiscent of a standard definition of a dimensionless universal renormalized coupling constant in ϕ^4 theory including the Ising limit with its infinite bare coupling. It is given by

$$g_R = -\frac{\chi_4}{(\chi_2)^2}m^D \tag{16}$$

where m is a the renormalized mass.

Often, for example in [2], [1], the mass is defined in terms of the two point function in an infinite volume at vanishing momentum. We substitute this by a definition using the two

smallest possible momenta in a periodic volume as in [8]. The two point function (11) in momentum space may be measured by

$$\tilde{G}(p) = \frac{\langle\langle \cos(p(u-v))\rangle\rangle}{\langle\langle \delta_{u,v}\rangle\rangle}. \tag{17}$$

Our definition of a renormalized mass m is

$$\frac{m^2}{m^2+\hat{p}_*^2} = \frac{\tilde{G}(p_*)}{\tilde{G}(0)} = \langle\langle \cos(p_*(u-v))\rangle\rangle \tag{18}$$

where we use the smallest momentum

$$p_* = (2\pi/L, 0, 0, \ldots, 0), \quad a\hat{p}_* = 2\sin(\pi a/L) \tag{19}$$

and average over its D possible directions. The rationale of the definition (16) is that it is a dimensionless ratio with the same number of fields in the numerator and denominator and hence it is expected to have a universal continuum limit. As it vanishes for Gaussian theories it is a measure of the interaction strength. We are thus led to the definition

$$g = 2\,\langle\langle \mathcal{X}(u,u',k+k')\rangle\rangle \times z^D, \quad z = mL. \tag{20}$$

Combining triviality with finite size scaling we investigate the proposition that the continuum limit at fixed z forces $g \searrow 0$. As for other questions on non-perturbative ultraviolet renormalization [10], [11], [12] we find it advantageous to employ a finite volume renormalization scheme also here. We shall perform a sequence of simulations of growing $L \equiv L/a$ where we tune β such as to maintain a fixed value $z = 2$. The advantage of keeping L finite and not too large in physical length units m^{-1} is that for the manageable values of L/a we expect to be closer to the universal continuum limit. If the theory is trivial we should find $g \to 0$ as $L/a \to \infty$. This is expected [3], [6], [4] for $D > 4$, and likely, although only at a logarithmic rate, for $D = 4$.

To get confidence in our algorithmic implementation we first reproduced the results of cluster simulations in [5] within errors. For the 20^4 lattice our error in g for a comparable number of Flops is about 12 times smaller.

In the table we compile our data. Each line corresponds to 10^6 iterations with V link updates in each of the two replica. The cost is dominated by the runs $D = 4, L = 32$ and $D = 5, L = 16$ with 240 hours each on a single PC. We refer to a code running under `matlab` but importing random numbers from the `ranlux` generator [13] in C [14]. We have insisted

D	L/a	β	z	$\mathcal{X}$	$\partial\mathcal{X}/\partial z$	$\mathcal{X}(z=2)$
4	8	0.148320	1.9981(27)	0.39235(96)	-0.3200(14)	0.39175(63)
4	10	0.148748	1.9949(26)	0.37256(92)	-0.3193(14)	0.37093(62)
4	12	0.148996	1.9992(26)	0.35493(91)	-0.3165(15)	0.35469(60)
4	16	0.149270	1.9988(25)	0.33161(91)	-0.3129(16)	0.33125(58)
4	22	0.149449	2.0085(24)	0.30831(86)	-0.3030(16)	0.31088(57)
4	32	0.149571	1.9956(24)	0.29028(83)	-0.2993(20)	0.28896(55)
3	8	0.217350	1.9946(36)	0.59387(100)	-0.2929(13)	0.59228(76)
3	10	0.218560	1.9942(37)	0.58634(102)	-0.2950(13)	0.58463(76)
3	16	0.220153	2.0047(37)	0.57240(109)	-0.3023(14)	0.57382(77)
3	32	0.221143	2.0032(39)	0.56338(118)	-0.3076(17)	0.56435(76)
5	8	0.113052	2.0041(18)	0.19126(59)	-0.2390(13)	0.19223(45)
5	10	0.113336	1.9993(16)	0.16037(53)	-0.2170(13)	0.16022(41)
5	12	0.113503	1.9937(15)	0.13884(47)	-0.1957(12)	0.13760(38)
5	16	0.113674	1.9918(13)	0.10944(39)	-0.1656(12)	0.10809(33)

TABLE I: Simulation results for $D = 3, 4, 5$.

on luxury level 2, but this part still accounts for only 4 % of the CPU time. The main code could clearly be accelerated substantially. Derivatives of $\mathcal{X}$ and z with respect to $\ln\beta$ can be measured as connected correlations with $\sum_l k(l)$ and their quotient yields our estimate for $\partial\mathcal{X}/\partial z$. In the end it emerges as a certain nonlinear function of primarily measured observables and it as well as all other errors is estimated by the tools provided in [15]. We take its measured value in each data set as a fixed constant and then form, now with $\mathcal{X}$ and z as functions of primary quantities, the combination

$$\mathcal{X}(z=2) = \mathcal{X} + (2-z)\partial\mathcal{X}/\partial z. \tag{21}$$

A look at the table shows that this is a tiny but sometimes significant correction. We can however safely neglect the error of $\partial\mathcal{X}/\partial z$ and higher terms of the Taylor expansion. It came as a pleasant surprise that even where no systematic correction is needed the statistical fluctuations in this combination partially cancel and thus reduce the error. This saves more than another factor two in run-time. The compensation is actually plausible: when sampled

graphs are 'bigger' than average $\mathcal{X}$ goes up but the mass goes down. In the table we note that relative errors are practically independent of L/a, which means complete absence of critical slowing down. Nevertheless the (short) autocorrelations do have to be taken into account when errors are determined.

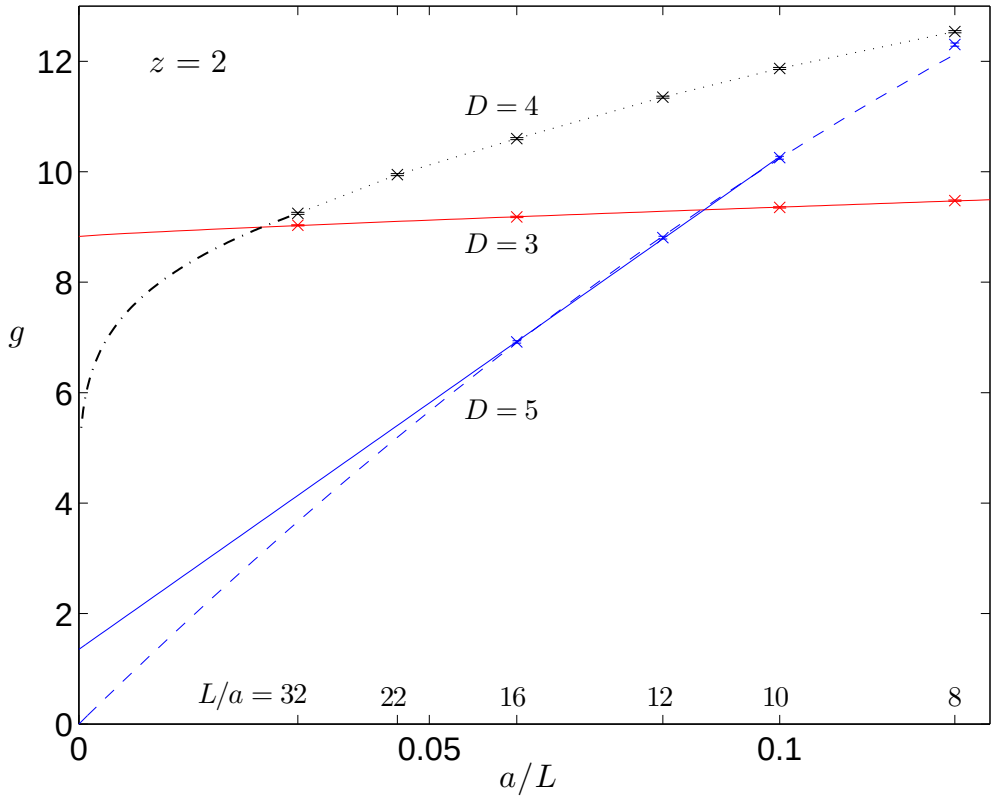

FIG. 1: All measured couplings with fits (full and dashed lines) discussed in the text. Errorbars are barely visible within the symbols.

In figure 1 all coupling data are plotted. Fits (full and dashed lines) all have acceptable χ^2. The $D = 3$ values are almost cutoff independent and approach a finite continuum value. The fit is $A + B(a/L)^\omega$ with the corrections to scaling exponent [16] $\omega = 0.85$. Due to the flatness, ω can vary over a wide range including $\omega = 1$. For $D = 5$ the full line is $A + Ba/L$ omitting the $L = 8$ lattice. The dashed fit is $Aa/L + B(a/L)^2$. A 32^5 simulation would be of interest to better verify triviality for this case. The $D = 4$ data show more curvature and the dotted lines just connect the data points. A naive 'eyeball' extrapolation in this plot against a/L would probably arrive at a nonzero value while on theoretical grounds we shall argue now for the dash-dotted line extrapolating to the origin at a vertical slope.

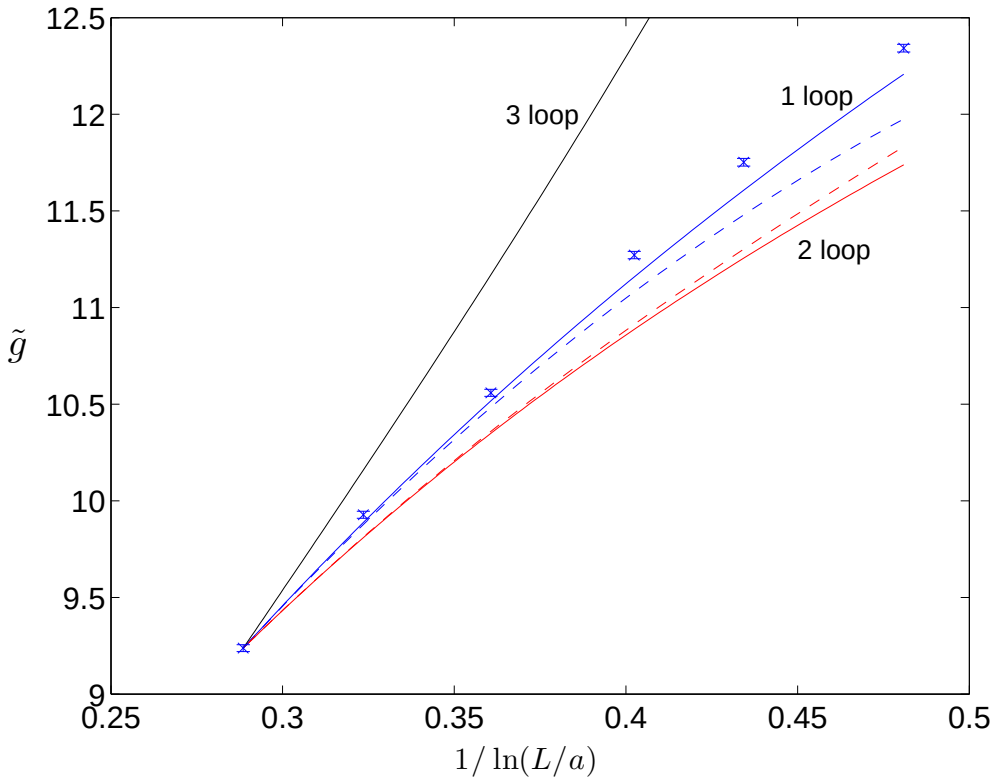

FIG. 2: Numerical couplings together with the renormalization group evolution starting from the finest lattice. Dashed lines include lattice artifacts of the β-function.

To confront the $4D$ data with the theory [2], [1] we plot against $[\ln(L/a)]^{-1}$ in figure 2. We solve the Callan Symanzik equation for the cutoff dependence of the coupling starting from $L = 32$. For this plot we have changed to the coupling

$$\tilde{g} = g[1 + (za/L)^2/8]^{-2} = g + \mathrm{O}(a^2) \tag{22}$$

differing by small cutoff effects only. For it in contrast to g there are no tree level artifacts in the β-function and also the cutoff corrections of the one and two loop terms are more uniform. We have worked out the lattice perturbation theory for our scheme up to two loops and could therefore, by relating to [1], also obtain the three loop term (without cutoff effects). Details will be reported elsewhere [17]. We here draw the following conclusions: The one loop result is accurate to a few percent for the scale changes considered. For instance, it accounts for 97% of the change $L/a = 32 \to 16$. The two loop term has a reasonable relative size but the wrong sign. The three loop term is the first one that is scheme dependent and hence depends on z. It is very large for $z = 2$ which suggests that renormalized perturbation

theory as an asymptotic expansion here fails to improve the leading order. It rather is at its limit with only the one loop approximation being numerically accurate at the percent level. Nonetheless it seems convincing to now trust the one loop approximation for $a < L/32$ (for $z = 2$) which implies a vanishing g in the continuum limit. On the way to it, also the higher loop terms should eventually cooperate to improve the approximation. The dash-dotted curve in figure 1 shows the one loop evolution continued. We plan a more detailed discussion of the perturbative series and data for other z values in [17].

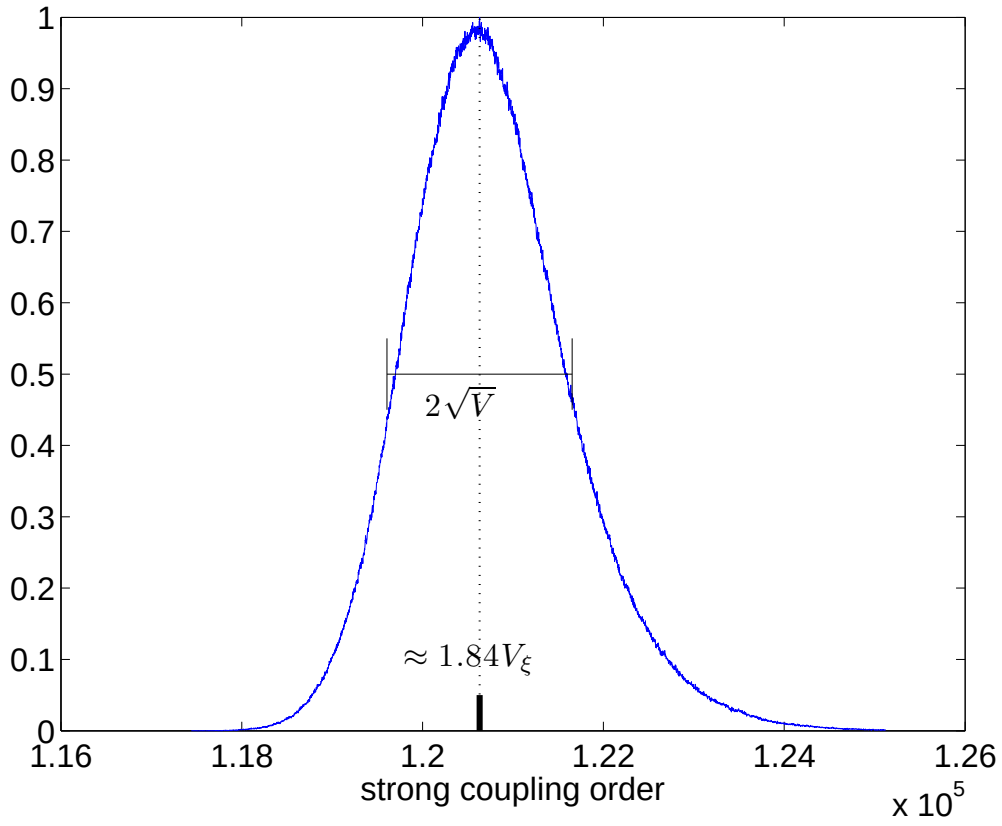

FIG. 3: The order in β of the diagrams sampled for $L = 32, D = 4$ ($u = v$ only).

We end on a more technical theme concerning the strong coupling simulation. The order of the diagrams that the algorithm has picked to be important for our physics is shown in figure 3. The peak is at the order of the correlation volume $V_\xi = (am)^{-4}$ and the width seems to be controlled by $(L/a)^2$.

Acknowledgments. Exchange with Peter Weisz, Erhard Seiler and Rainer Sommer was most beneficial to this project. Some part of the effort was stimulated during a visit at the Max Planck (Werner Heisenberg) Institut in München by hospitality and support. I

owe thanks to Willi Rath for helping me with `ranlux` under `matlab` and to the HU physics compute team for providing a smooth infrastructure. Finally financial support of the DFG via SFB transregio 9 is acknowledged.

---

[1] M. Lüscher, P. Weisz, Nucl. Phys. B290 (1987) 25.

[2] E. Brezin, J. C. Le Guillou, J. Zinn-Justin, Field Theoretical Approach to Critical Phenomena in *Phase Transitions and Critical Phenomena*, Vol.6, London 1976, 125.

[3] M. Aizenman, Phys. Rev. Lett. 47 (1981) 1.

[4] J. Fröhlich, Nucl. Phys. B200 (1982) 281.

[5] I. Montvay, G. Münster, U. Wolff, Nucl. Phys. B305 (1988) 143.

[6] M. Aizenman, Commun. Math. Phys. 86 (1982) 1.

[7] N. Prokof'ev, B. Svistunov, Phys. Rev. Lett. 87 (2001) 160601.

[8] Y. Deng, T. M. Garoni, A. D. Sokal, Phys. Rev. Lett. 99 (2007) 110601.

[9] U. Wolff, Nucl. Phys. B810 (2009) 491.

[10] U. Wolff, Nucl. Phys. B265 (1986) 537.

[11] M. Lüscher, P. Weisz, U. Wolff, Nucl. Phys. B359 (1991) 221.

[12] M. Lüscher, R. Narayanan, P. Weisz, U. Wolff, Nucl. Phys. B384 (1992) 168.

[13] M. Lüscher, Comput. Phys. Commun. 79 (1994) 100.

[14] M. Lüscher, http://luscher.web.cern.ch/luscher/ranlux.

[15] U. Wolff, Comput. Phys. Commun. 156 (2004) 143.

[16] M. Hasenbusch, J. Phys. A32 (1999) 4851.

[17] U. Wolff, in preparation.